



Accurate crystal fields for embedded cluster calculations

M. Klintenberg*, S.E. Derenzo, M.J. Weber

Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94 720, USA

Received 19 October 1999

Abstract

Local electronic structure defects in ionic crystals is commonly modeled using embedded cluster calculations. In this context we describe how to embed the quantum cluster (QC) in an array of point charges. Specifically, the method calculates an array of point charges that reproduces the electrostatic potential of the infinite crystal within an accuracy usually $< 1 \mu\text{V}$ in the interior of the QC. © 2000 Published by Elsevier Science B.V. All rights reserved.

PACS: 71.15.Th; 02.70.-c; 31.15.-p

Keywords: Embedded cluster calculation; Crystal field potential; Ewald potential

PROGRAM SUMMARY

Title of program: Ewald

Catalogue identifier: ADME

Program Summary URL: <http://cpc.cs.qub.ac.uk/summaries/ADME>

Computers: Silicon Graphics (Origin 2000 and Power challenge)

Operating systems or monitors under which the program has been tested: IRIX 6.2, 6.4 and 6.5.5

Programming language used: C (cc 7.0 for IRIX 6.2 and 6.4, cc 7.3 for IRIX 6.5.5)

Memory required to execute with typical data: 16 Mwords

Number of bits in a word: 32

Number of processors used: 1

Has the code been vectorized or parallelized?: No

Number of bytes in distributed program, including test data, etc.: 19 367 bytes

Distribution format: tar gzip file

Keywords: Embedded cluster calculation, crystal field potential, Ewald potential

Nature of physical problem

The electrostatic potential, due to the crystal field (CF) in a periodic system can be calculated using the Ewald summation method [1]. To model local electronic structure defects in ionic crystals a periodic approach should not be used because a periodic approach can not handle defects with charge or dipole moment. Instead the system is modeled as a quantum cluster (QC) embedded in an array of point charges. In this paper we describe how to chose these point

* Corresponding author: Lawrence Berkeley National Laboratory, One Cyclotron Road #55-121, University of California, Berkeley, CA 94720, USA; Fax: (510) 486-4768; E-mail: mkklintenberg@lbl.gov.

charges to obtain Ewald-like electrostatic potentials in the interior of the QC.

Method of solution

The method of solution can be divided into 5 steps. (1) An array of typically $1.0e4$ point charges is generated using $2N \times 2N \times 2N$ unit cells. (2) Three zones are identified. Zone 1 is defined by the QC atoms. Zone 2 consists of approximately 500 point charges that embed the QC spherically. These point charges have their ionic charges. Zone 3 is the remainder of the array. These point charge values are later treated as parameters. (3) The Ewald potentials are calculated for all sites in zone 1 and 2. (4a) If the system has any dipole moment this is removed using a random Δq approach. (Note that this dipole moment could be removed directly in the system of linear equations but we have found that better solutions are usually obtained if the dipole moment is removed before solving the system of equations.) (4b) Solving a system of linear equations give the charge values (Δq 's) for the zone 3 sites so that the direct sum electrostatic potential in zone 1 and 2 become equal to the corresponding Ewald values. (5) Using 1000 randomly chosen points in the interior of the QC the result is checked and the Δ_{rms} is computed.

$$\Delta_{\text{rms}} = \left(\sum_i^{N_r} [V_{\text{Ewald}}(\mathbf{r}_i) - V_{\text{DS}}(\mathbf{r}_i)]^2 / N_r \right)^{1/2}, \quad (1)$$

N_r is the number of random points in the QC, $V_{\text{Ewald}}(\mathbf{r}_i)$ and $V_{\text{DS}}(\mathbf{r}_i)$ are the Ewald and direct sum potential at \mathbf{r}_i , respectively. Δ_{rms} is usually $< 1 \mu\text{V}$ using our algorithm.

Restrictions on the complexity of the program

The algorithm can be applied to any ionic crystal where the unit cell information is available.

Typical running time

Typical running time is less than 40 minutes on a R10000 processor, slightly longer on a PII 450 MHz (Linux) system.

Unusual features of the program

The numerical libraries CLAPACK and CBLAS [2] are used for the linear algebra operations. dgelsx computes the minimum norm least squares solution to an over- or under-determined system of linear equations using a complete orthogonal factorization. A pseudo-random number generator (randome.c) [3] is also provided and needed for the code.

References

- [1] P.P. Ewald, Ann. Phys. 64 (1921) 253.
- [2] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, D. Sorensen, LAPACK Users' Guide (Society for Industrial and Applied Mathematics, Philadelphia, PA, 1995) ISBN 0-89871-345-5. <http://www.netlib.org/clapack/>.
- [3] P. L'Ecuyer, Comm. ACM 31 (1988) 742–749.

LONG WRITE-UP

1. Theoretical background

Defect studies in inorganic crystals using modern quantum chemistry codes usually require embedding of a carefully chosen quantum cluster of atoms in an array of point charges. A periodic approach, i.e. using periodic boundary conditions (PBC), is of course tempting but difficult to implement because by definition defect studies are non-periodic and charged defects are difficult to handle. In this context we have developed a novel algorithm for choosing the point charge array used in the embedding procedure.

Naively, an array of unit cells can be placed around the QC. However, in most cases this approach results in large deviations from the Ewald potential. As an example consider CaF_2 . Placing $120 \times 120 \times 120$ unit cells (approximately $2.1e7$ ions) around a QC give an rms deviation between the Ewald potential and the direct sum potential of $\Delta_{\text{rms}} = 17.281 \text{ eV}$!

The algorithm we describe in this paper sets the charge of the ions in the point charge array (conserving total charge and reducing any dipole moment to zero) so that the direct sum potential reproduces the Ewald potential with an rms accuracy $\Delta_{\text{rms}} < 1 \mu\text{V}$. This value is obtained by choosing $1e3$ random points in the interior of the QC, computing both the Ewald- and direct-sum potentials, then calculating the rms deviation (Δ_{rms}).

A number of techniques have been developed to tackle the problems that occur when dealing with long range interactions such as the Coulomb interaction, see Refs. [1–3] for details. Because sums of the type

$$V(\mathbf{r}_i) = \sum_{\mathbf{n}} \sum_{j=1}^N \frac{q_j}{r_{ij,\mathbf{n}}} \quad (2)$$

are slowly and conditionally convergent, great precaution is needed. Ionic crystals are best dealt with using lattice methods, such as the Ewald summation method. This approach is a method for summing the interaction between an ion and its periodic images [2]. Because of the conditional convergence of Eq. (2) the final result will depend on in which order the periodic images are summed. The most natural summation scheme is of course to add the unit cells spherically. Depending on the dielectric constant (relative permittivity) of the medium that surrounds the very large “spherical” set of unit cells, the summation result will come out differently. In the case of vacuum, a dipole layer (due to the dielectric constant of the surrounding medium, the total dipole moment of the unit cell and the shape of the macroscopic lattice) will develop and has to be compensated for [4], and in the case of a perfect conductor ($\epsilon_s = \infty$) the dipole layer will not develop. The Ewald approach is an efficient method for dealing with sums such as Eq. (2) in the limit of a very large sphere with a perfect conductor surrounding. Details on the Ewald technique and the physical interpretation of the different terms can be found in Refs. [1–3].

The Ewald potential is given by [5]

$$V(\mathbf{r}_i) = \sum_j \sum_{\mathbf{n}} q_j \frac{\text{erfc}(\eta r_{ij,\mathbf{n}})}{r_{ij,\mathbf{n}}} + \frac{1}{\pi V} \sum_j q_j \sum_{\mathbf{m} \neq 0} \frac{\exp(-[\pi f_{\mathbf{m}}/\eta]^2)}{f_{\mathbf{m}}^2} \times \cos(2\pi \mathbf{f}_{\mathbf{m}} \cdot \mathbf{r}_{ij,0}). \quad (3)$$

$r_{ij,\mathbf{n}}$ is given by

$$r_{ij,\mathbf{n}} = |\mathbf{r}_i - \mathbf{r}_{j,\mathbf{n}}|, \quad \mathbf{r}_{j,\mathbf{n}} = (u_{1,j} + n_1)\mathbf{a}_1 + (u_{2,j} + n_2)\mathbf{a}_2 + (u_{3,j} + n_3)\mathbf{a}_3.$$

$\mathbf{r}_{j,\mathbf{n}}$ is the Cartesian coordinates for ion j in unit cell (n_1, n_2, n_3) and $(u_{j,1}, u_{j,2}, u_{j,3})$ is the fractional unit cell coordinate for ion j . $\mathbf{a}_k = (a_{k,x}, a_{k,y}, a_{k,z})$, $k = 1, 2, 3$ are the three unit cell translation vectors. The reciprocal lattice space vectors are given by

$$\mathbf{b}_k = (b_{k,1}, b_{k,2}, b_{k,3}), \quad k = 1, 2, 3$$

and

$$\mathbf{b}_1 = \mathbf{a}_2 \times \mathbf{a}_3 / V, \quad \mathbf{b}_2 = \mathbf{a}_3 \times \mathbf{a}_1 / V, \quad \mathbf{b}_3 = \mathbf{a}_1 \times \mathbf{a}_2 / V.$$

The volume is of course $V = \mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)$. $\mathbf{f}_{\mathbf{m}}$ is the position vector in reciprocal space

$$\mathbf{f}_{\mathbf{m}} = (m_1, m_2, m_3) \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}. \quad (4)$$

The direct sum potential is simply

$$V(\mathbf{r}_i) = \sum_{j=1}^N \frac{q_j}{r_{ij}}. \quad (5)$$

Note that N now is the total number of atoms considered and not the number of atoms in the unit cell as was the case above. Fig. 1 shows a schematic picture of the QC and the point charge array.

Zone 1 is defined as the QC. Zone 2 consists of approximately 500 point charges where the ionic charge values are not altered. These ions all have their formal ionic charges. Zone 3 is defined by the rest of the ions in the system and usually consists of around $1e4$ charges whose values are treated as parameters in the fitting procedure. The

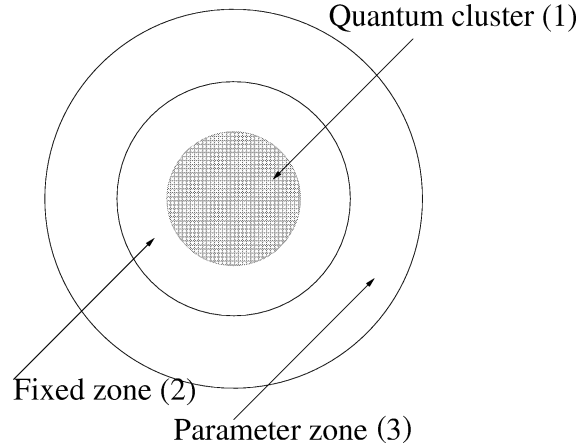


Fig. 1. Schematic picture of the QC and the point charge array.

number of zones 2 and 3 atoms should be considered as parameters of the method (to our knowledge not related to physical properties such as polarizability of the ionic crystal). However, it was noted that for large unit cells the number of zones 2 and 3 atoms had to be larger than what was needed for small unit cells to obtain comparable accuracy. Of course these numbers can be optimized for each ionic crystal if necessary for numerical reasons or computational limitations.

The system of linear equations is defined in Δq_k because the charge changes must be small. The i th equation is given by

$$\sum_k^N \frac{q_k + \Delta q_k}{r_{ik}} - V_{\text{Ewald}}(\mathbf{r}_i) = 0, \quad (6)$$

$$N = N_1 + N_2 + N_3, \quad (7)$$

$$\Delta q_k = 0, \quad k \leq N_1 + N_2, \quad (8)$$

$$q_k \in \begin{cases} \text{zone 1} & 1 \leq k \leq N_1, \\ \text{zone 2} & N_1 < k \leq N_2, \\ \text{zone 3} & N_2 < k \leq N \end{cases} \quad (9)$$

and the four extra equations (conservation of charge and dipole moment) are given by

$$\sum_k^N \Delta q_k = 0, \quad \text{and} \quad \sum_k^N \Delta q_k r_{kl} = 0, \quad l = x, y, z. \quad (10)$$

As expected the Δq values are small. Note that the dipole moment (if any) is driven to zero before solving the system of linear equations. This is done by randomly picking zone 3 charge pairs and adding $+\Delta q$ and $-\Delta q$, respectively. This procedure is repeated until a threshold value for the dipole moment is reached.

2. Program structure and comments

The program structure is straightforward and easily understood. First input values are read, necessary storage is allocated and various initiations are made. Using the unit cell information all geometrical structures are built, e.g., real and reciprocal space vectors. The checkpoints in the cluster are generated and initial potential values are

computed. The Ewald potentials are then calculated. Note that the Ewald convergence parameter $\eta = 0.2$ as well as the sum limits of $\mathbf{n}_{\max} = (8, 8, 8)$ and $\mathbf{m}_{\max} = (5, 5, 5)$ are set in the code. We have so far not encountered any problems with the values chosen. If any dipole moment is present it is removed by the procedure described above. Finally the system of linear equations is solved, the Δ_{rms} is computed, and file outputs are written.

2.1. Installation of the program

Uncompress and tar the program where the Ewald directory shall live

```
$uncompress ewald.tar.Z
$tar xvf ewald.tar
```

For installation of the numerical libraries (CLAPACK and CBLAS), see instructions given in [6].

2.2. Input files

The following input files are needed: ewald.in, name.uc, name.qc and seedfile. ewald.in has the following format:

```
name
nr of random constraint points in zone 1
nr of constraint atoms defining sphere containing zones 1 and 2
nr of atoms defining zone 1 0 → read from .rmspot file
```

The name.uc file has the format:

```
a1x a1y a1z Nx
a2x a2y a2z Ny
a3x a3y a3z Nz

fx fy fz q type
...
```

The total number of ions in the run ($N_T = N_1 + N_2 + N_3$) is $2N_x \times 2N_y \times 2N_z \times$ number of ions per unit cell. fx, fy and fz are the fractional unit cell coordinates.

The name.qc contains the quantum cluster. The coordinates in this file has to be consistent with the name.uc file. name.qc has the following format:

```
type x y z q
...
```

The seedfile has to be created only once in each directory where the program is run. The file consist of two numbers:

```
seed1 seed2
```

Seed1 should be chosen in the range $[1, 2^{31} - 86]$ and seed2 in the range $[1, 2^{31} - 250]$. In the *Selected examples* section below explicit examples of in- and output files will be shown.

2.3. Output files

The program generates the following output files: name.out, name.clust, name.pts-nw, name.pts-jag, seedfile and tmpfile.

The name.out file contains all of the output from the program. The amount of information printed is set by the DEBUG level (an intrinsic parameter which has to be set in the code, default DEBUG = 0 generates the least amount of output). The very tail of this file contains the rms results, charge ranges, etc.

name.clust holds the cluster definition. A spherical cluster is generated adding one spherical shell per iteration until \geq nrclusteratoms ions are counted. Note that the nrclusteratoms_rmax_rnd variable is only important for the Δ_{rms} check. This variable defines the region within which the Δ_{rms} is calculated. If nrclusteratoms_rmax_rnd is = 0, the Δ_{rms} volume is defined by name.qc, and if the variable is set \neq 0, the Δ_{rms} volume is defined by at least nrclusteratoms_rmax_rnd ions (spherical shells are added as above).

name.pts-jag and name.pts-nw are two input files generated for JAGUAR [7] and NWChem [8], respectively. These files have cluster- and point-charge definitions. Of course the files need further editing to run with the two quantum chemistry codes specified, but the back bone is defined. If nrclusteratoms_rmax_rnd = 0, the cluster printed in these files are defined by name.qc, and if nrclusteratoms_rmax_rnd \geq 0 we have situation as above with at least nrclusteratoms_rmax_rnd ions.

The seedfile is rewritten automatically each run to guarantee new random points in each run (see the randome.c source for more information).

The tmpfile file is a scratch file used for temporary output.

3. Selected examples

In the top directory of the distribution tree the testing/ directory is found. By copying the content of this directory

```
$cp testing/* .
```

and compiling

```
$gcc -o Ewald Ewald.c randome.c -llapack -lblas -lF77 -lm
```

the following two tests can be run. The CaF₂ run is started by typing

```
$cp ewald.in.CaF2 ewald.in
$cp CaF2.seed seedfile
$Ewald < ewald.in
```

and the Lu₂SiO₅ test is run by typing

```
$cp ewald.in.Lu2SiO5 ewald.in
$cp Lu2SiO5.seed seedfile
$Ewald < ewald.in
```

The results are presented below.

3.1. Ca₅F₁₀

In this example a cluster is read from Ca5F10.666.qc, the point charge array is $12 \times 12 \times 12$ ($N = 6$ which gives 20 736 sites), there are 1000 random test points within the cluster region and zone 1 and 2 add up to at least 500 sites (the actual value turns out to be 537). The ewald.in file looks as follows:

```
Ca5F10.666
1000 Enter number of random check points in zone 1:
500 Enter nr constraint atoms def. sphere containing zones 1 and 2:
```

0 Enter nr atoms defining zone 1 sphere: 0 → read from .qc file

The Ca5F10.666.uc is:

```
5.453 0.0 0.0 6
0.0 5.453 0.0 6
0.0 0.0 5.453 6
0.75 0.5 0.75 2.0 Ca
...
0.0 0.25 0.0 -1.0 F
```

The Ca5F10.666.qc is:

```
Ca 0.00000 0.00000 0.00000 2.0
...
F 4.08975 1.36325 1.36325 -1.0
```

The tail of Ca5F10.666.out will look:

```
imad 1535 x -0.57140 y -1.80806 z -1.93113 q 0.0000 ????
      dspot 9.166142 V ewald -0.022803 V F 9.188945 V
imad 1536 x 3.07121 y -3.30131 z -0.70990 q 0.0000 ????
      dspot 4.724824 V ewald 9.255035 V F -4.530211 V
```

Results before solving the system of linear eq.

```
(Ewald - direct sum) potential average 8.292521 rms 38.580914
dipole dx -10.906000 dy -10.906000 dz -10.906000 d2 356.822508
radmom 77.254339
unit cell: energy -62.692808
```

lapack...

Results after solving the system of linear eq.

```
q ranges: -1.343521 to 2.306819
rmspot = 0.000000
```

Not only the random chk points: rmspot=1.560674e-08

Only the random chk points: rmspot=1.934858e-08

Unit cell has 4 atoms of type Ca

Unit cell has 8 atoms of type F

This gives Z=4

Madelung constant 5.038801 for R(Ca-F) 2.361218

seed1 = 1151145213, seed2 = 1057074022

Started: Ca5F10.666 Wed Oct 6 03:32:30 1999

Finished: Ca5F10.666 Wed Oct 6 04:05:04 1999

END

The seedfile started with these two numbers 911 513 430 and 967 139 521

3.2. Lu₂SiO₅ (LSO)

In this example a cluster is not read from Lu2SiO5.333.qc, the point charge array is $6 \times 6 \times 6$ ($N = 3$ which gives 13 824 sites), there are 1000 random test points within the cluster region and zone 1 and 2 add up to at least 500 sites (the actual value turns out to be 500). The ewald.in file looks as follows:

Lu2SiO5.333

```
1000 Enter number of random check points in zone 1:
500 Enter nr constraint atoms def. sphere containing zones 1 and 2:
50 Enter nr atoms defining zone 1 sphere: 0 → read from .qc file
```

The Lu2SiO5.333.uc is:

```
14.2770 0.0000 0.0000 3
0.0000 6.6398 0.0000 3
-5.46374818 0.0000 8.66823039 3
0.5371 0.7563 0.4664 3.0 Lu1
...
0.5175 0.0980 0.3990 -2.0 O5
```

The Lu2SiO5.333.qc is not necessary because nrclusteratoms_rmax_rnd=50.

The tail of Lu2SiO5.333.out will look:

```
imad 1498 x 5.42978 y 1.19108 z -0.31504 q 0.0000 ????
dspot -3.772371 V ewald -3.934661 V F 0.162290 V
imad 1499 x -4.53610 y -3.37202 z -0.44519 q 0.0000 ????
dspot -11.246114 V ewald -11.429500 V F 0.183386 V
```

Results before solving the system of linear eq.

```
(Ewald - direct sum) potential average -0.173901 rms 0.473479
dipole dx 0.000000 dy 0.000000 dz 0.000000
d2 0.000000 radmom 135.853038
unit cell: energy -2145.932110
```

lapack...

Results after solving the system of linear eq.

```
q ranges: -2.027005 to 4.009840
rmspot = 0.000000
```

```
Not only the random chk points: rmspot=1.016383e-07
Only the random chk points: rmspot=1.244809e-07
```

```
Unit cell has 8 atoms of type Lu1
Unit cell has 8 atoms of type Lu2
Unit cell has 8 atoms of type Si
Unit cell has 8 atoms of type O1
Unit cell has 8 atoms of type O2
Unit cell has 8 atoms of type O3
```



```
Unit cell has 8 atoms of type 04  
Unit cell has 8 atoms of type 05  
This gives Z=8
```

```
Madelung constant 35.491122 for R(Si-O3) 1.604038
```

```
seed1 = 56775299, seed2 = 789623131
```

```
Started: Lu2SiO5.333 Wed Oct 6 02:43:23 1999  
Finished: Lu2SiO5.333 Wed Oct 6 03:18:10 1999  
END
```

The seedfile started with the following two numbers 1 151 145 213 and 1 057 074 022.

4. Concluding remarks

The algorithm presented is a straightforward implementation for choosing point charge arrays that properly embed an arbitrary shaped atom cluster. The electrostatic potentials in the interior of the cluster (at arbitrary points) agree with those computed using the Ewald summation method. The Δ_{rms} is usually less than 1 μV which is well within the accuracy needed for quantum cluster calculations.

Acknowledgments

This work was supported in part by the Director, Office of Energy Research, Office of Biological and Environmental Research, Office of Medical Sciences, of the U.S. Department of Energy under contract No. DE-AC03-76SF00098, and in part by the National Institutes of Health, National Cancer Institute under grant No. R01-CA48002, and in part by The Swedish Natural Science Research Council.

References

- [1] A.Y. Toukmaji, J.A. Board Jr., *Comput. Phys. Commun.* 95 (1996) 73.
- [2] M. Allen, D. Tildesley, *Computer Simulation of Liquids* (Oxford Science, London, 1990).
- [3] C. Kittel, *Introduction to Solid State Physics* (John Wiley and Sons, New York, 1971).
- [4] S.W. De Leeuw, J.W. Perram, E.R. Smith, *Proc. Roy. Soc. London A* 373 (1980) 27.
- [5] P.P. Ewald, *Ann. Phys.* 64 (1921) 253.
- [6] <http://www.netlib.org/clapack/>.
- [7] Jaguar v3.5 (Schrödinger, Inc., Portland, OR, 1997).
- [8] NWChem v3.3 (PNNL, Richland, WA, 1998).